# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No. 0704-0188*



| 1. REPORT DATE *(DD-MM-YYYY)* | 2. REPORT TYPE | 3. DATES COVERED *(From - To)* |
|---|---|---|
| 05-18-2015 | FINAL | |

**4. TITLE AND SUBTITLE**

Strong Artificial Intelligence and National Security

**5a. CONTRACT NUMBER**

**5b. GRANT NUMBER**

**5c. PROGRAM ELEMENT NUMBER**

**6. AUTHOR(S)**

Jonas Stewart

Paper Advisor (if Any): Associate Professor Christopher J. McMahon

**5d. PROJECT NUMBER**

**5e. TASK NUMBER**

**5f. WORK UNIT NUMBER**

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**

Joint Military Operations Department
Naval War College
686 Cushing Road
Newport, RI 02841-1207

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)**

**10. SPONSOR/MONITOR'S ACRONYM(S)**

**11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

**12. DISTRIBUTION / AVAILABILITY STATEMENT**

*Distribution Statement A: Approved for public release; Distribution is unlimited.*
*Reference: DOD Directive 5230.24*

**13. SUPPLEMENTARY NOTES** A paper submitted to the Naval War College faculty in partial satisfaction of the requirements of the Joint Military Operations Department. The contents of this paper reflect my own personal views and are not necessarily endorsed by the NWC or the Department of the Navy.

**14. ABSTRACT**

**Prominent business and science leaders believe that technological advances will soon allow humankind to develop artificial intelligence (AI) that meets and even exceeds human cognitive capabilities. This "strong" AI will potentially revolutionize national security affairs by decreasing the human cost of war while increasing the speed and efficiency of America's tools of national power at the tactical, operational, and strategic levels of conflict. However, with these benefits comes the risk of a future AI race, or even the advent of a superintelligent, hostile AI. As national security experts, including combatant commanders, prepare for the arrival of strong AI, they must develop methods to confront enemy AI with ill-structured problems in the battlefield in order to defeat its faster, yet more linear thought processes. On the strategic level, American national security experts must also press for both national and international regulation of this new technology to harness its tangible operational benefits while avoiding its potential strategic pitfalls.**

**15. SUBJECT TERMS**

Artificial Intelligence

| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT | 18. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON |
|---|---|---|---|---|---|
| a. REPORT | b. ABSTRACT | c. THIS PAGE | | | Chairman, JMO Dept |
| UNCLASSIFIED | UNCLASSIFIED | UNCLASSIFIED | | 26 | 19b. TELEPHONE NUMBER *(include area code)* 401-841-3556 |

**Standard Form 298 (Rev. 8-98)**

**NAVAL WAR COLLEGE**
**Newport, R.I.**

# Strong Artificial Intelligence and National Security: Operational and Strategic Implications

**by**

**Jonas Stewart**

**U.S. Department of State**

**A paper submitted to the Faculty of the Naval War College in partial satisfaction of the requirements of the Department of Joint Military Operations.**

**The contents of this paper reflect my own personal views and are not necessarily endorsed by the Naval War College, the Department of the Navy, or the Department of State.**

**Signature: _____**

**18 May 2015**

# Contents

## Paper Abstract

*Strong Artificial Intelligence: Operational and Strategic Implications*

Prominent business and science leaders believe that technological advances will soon allow humankind to develop artificial intelligence (AI) that meets and even exceeds human cognitive capabilities. This "strong" AI will potentially revolutionize national security affairs by decreasing the human cost of war while increasing the speed and efficiency of America's tools of national power at the tactical, operational, and strategic levels of conflict. However, with these benefits comes the risk of a future AI race, or even the advent of a superintelligent, hostile AI. As national security experts, including combatant commanders, prepare for the arrival of strong AI, they must develop methods to confront enemy AI with ill-structured problems in the battlefield in order to defeat its faster, yet more linear thought processes. On the strategic level, American national security experts must also press for both national and international regulation of this new technology to harness its tangible operational benefits while avoiding its potential strategic pitfalls.

# Introduction

In January 2015, the Future of Life Institute (FLI), a conglomerate of senior business and science leaders, including Tesla/Space X founder Elon Musk and theoretical physicist Stephen Hawking, released an open letter warning of the existential risk presented by the next phase of artificial intelligence (AI) research on the future of humanity. They contend that market forces will soon accelerate research to the point where AI systems may reach and possibly exceed human capabilities. As a result, this "strong" AI, will soon transition from a novel scientific concept to an economically viable technology for real world applications. The letter concludes that humankind must consider and prepare for the ramifications of introducing this revolutionary technology into the world.[1]

Members of the national security community may balk at such grand warnings. After all, human history is riddled with technophobes who have warned of the dangers of the next great technology on the world. For example, during the early days of rail travel, many feared the human body could not withstand speeds exceeding 15 miles per hour.[2] Like luddites of the past, writers of science fiction novels and movies have warned us for decades of the impending threat of strong AI.[3]

This paper contends that, given both the potential risk involved, and the growing number of voices among our business and science elites behind the FLI's warning, the advent

[1] The Future of Life Institute, 2015, "Research Priorities for Robust and Beneficial Artificial Intelligence: an Open Letter," *The Future of Life Institute,* January 5, Accessed March 1, 2015. http://futureoflife.org/misc/open_letter.

[2] We now know that it can withstand speeds beyond Mach 2. How Stuff Works, n.d., "Old Railroads," *How Stuff Works,* Accessed April 21, 2015, http://history howstuffworks.com/american-history/old-railroads htm.

[3] The Luddites were a group of 19th-century English textile workers that protested the use of technology to economize manufacturing processes. The term is now a popular label for individuals or groups that oppose new technologies.

of strong AI deserves the U.S. national security establishment's attention. That premise then begs the question: How should senior national security officials, including combatant commanders, prepare for the arrival of strong AI and its impact on national security affairs? This paper contends that, on the tactical and operational levels, commanders should prepare to counter enemy AI on the battlefield with ill-structured problems in order to challenge its superior, but probably linear, decision-making processes. On the strategic level, senior national security leaders should establish national and international regulations governing the development of strong AI and its use in national security affairs. A brief survey of the history of AI development, the impact of strong AI on national security, and potential risks involved with the introduction of this new technology informs this conclusion.

## A Brief History of AI

To understand the potential impact of strong AI on national security, it is important to first define it. For the purpose of this paper, AI is a sub-discipline of Cognitive Science; an interdisciplinary field of study that examines the mind and intelligence.[4] AI is the study of how to create artificial systems that "think." These systems, known as "artificial agents," should be able to sense environmental variables, analyze them, and then make the best possible decision taking those variables into account.[5]

[4] Nayef R.F. Al-Rodhan, 2011, *The Politics of Emerging Strategic Technologies: Implications for Geopolitics, Human Enhancement and Human Destiny,* London: Palgrave Macmillan, 152.
[5] David Poole, Alan Mackworth, and Randy Goebel, 1998, *Computational Intelligence: A Logical Approach,* New York: Oxford University Press, Preface.

## History of AI Research

While humankind has long dreamed about using technology to create sentient beings, it was not until World War II that rudimentary computing technologies made AI development possible. Early computing pioneers, including British codebreaker Alan Turing, theorized that artificial systems could replicate thought processes.[6] AI became an official field when American computer scientist John McCarthy and others organized a summer research project on it at Dartmouth College in 1956.[7] Around that time, the U.S. government funded several AI research projects, at the RAND Corporation and elsewhere, with the purpose of developing assorted defense-related systems.[8]

Research and interest in AI have since gone through several cycles of promise and disappointment, known as "AI winters." Jim Howe from the University of Edinburgh writes that the field experienced its first winter when Sir James Lighthill, an applied mathematician, published a report in 1973 that questioned AI's fundamental capability to solve problems in the real world due to the almost infinite number of variables involved. Howe contends that the Lighthill report caused a significant drop in interest and funding for AI projects (including DARPA funding for defense-specific projects).[9] Following a surge of AI interest and funding in Japan and the United States in the 1980s, another AI winter occurred in the

---

[6] Noel Sharkey, 2012, "Alan Turing: The experiment that shaped artificial intelligence," *BBC News,* June 21, Accessed April 11, 2015, http://www.bbc.com/news/technology-18475646; Alan M. Turing, 1950, "Computing Machinery and Intelligence," *Mind,* 433-460, http://loebner.net/Prizef/TuringArticle.html.

[7] John McCarthy, M. L. Minsky, N. Rochester, and C. E. Shannon, 1955, "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence," *Stanford University: John McCarthy's Home Page,* August 31, Accessed April 11, 2015, http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth html.

[8] Philip Klahr, and A. Waterman Donald, 1986, *Artificial Intelligence: A RAND Perspective,* Report, RAND.

[9] DARPA is the "Defense Advanced Research Projects Agency." Founded in response to the surprise Soviet 1957 launch of Sputnik, DARPA pursues the use of emerging technologies to promote U.S. national security; Jim Howe, 2007, "Artificial Intelligence at Edinburgh University: A Perspective," *University of Edinburgh School of Informatics,* Accessed April 11, 2015, http://www.inf.ed.ac.uk/about/AIhistory html.

early 1990s when a major DARPA-funded project, the Strategic Computing Initiative, switched focus from AI to supercomputing research.[10]

It appears as though a third wave of AI hype has begun. Writing for the *Financial Times*, Richard Waters reports that investors poured over half a billion dollars into AI research during the latter half of 2014. He adds that private capital is flooding the AI sector now because developers think computing hardware, the major obstacle for previous attempts to develop strong AI, is approaching adequate levels. Writing for *Wired*, Sentient Technologies founder Babak Hodjat points to recent breakthroughs in machine learning, expanding access to AI development software, and increasing public awareness of advanced AI research as major drivers behind this third era of AI hype.

## Categories of AI

As mentioned earlier, rudimentary AI systems that engage in basic thinking functions already exist. In "The Politics of Emerging Strategic Technologies," Nayef Al-Rodhan writes that these current systems break down into Symbolic, Connectionist, and Evolutionary AI. Symbolic AI refers to agents that apply rule systems that attempt to replicate human decision-making processes. These agents are adept at calculating complex problems using symbols (like numbers, key concepts, etc…), but their "thinking" is limited by the sheer lines of programing required to catch the nuances of the human thought process. As a result, Symbolic AI's weakness is its lack of implicit (or common sense) knowledge.[11]

---

[10] Adam Elkus, 2015, "The A.I. Wars? Why Artificial Intelligence may not Revolutionize Security and Geopolitics-Yet," *Slate.com*, January 20, Accessed March 20, 2015, http://www.slate.com/articles/technology/future_tense/2015/01/what_artificial_intelligence_does_and_does_not_mean_for_security_and_geopolitics.html.

[11] Al-Rodhan, *The Politics of Emerging Strategic Technologies*, 155-156.

Ray Kurzweil writes about Connectionist AI in his book, *The Singularity is Near*. Based more on neuroscience than symbolic logic systems, Connectionist AI agents use artificial neuron networks to replicate the human mind at the organic level. The major advantage of Connectionist AI is that it can learn new information by testing existing hypotheses. While these agents learn from experience, they still lack some of the implicit knowledge required to approach human-level intelligence.[12]

Evolutionary AI involves several different AI systems that are placed in an evolutionary environment where they compete for end use. This environment is designed to replicate mate selection and other biological evolutionary processes. Al-Rodhan writes that NASA scientists have already used this emerging field of AI to develop space communications technologies by pitting various communications programs against each other until the most robust version remained.[13]

Experts contend that these fields of AI, combined with new programming process, including the ability to map the human brain using advanced scanning technologies, will soon allow humankind to develop AI agents that reach and exceed human capabilities. Now that strong AI has entered the realm of the possible, it is important to consider AIs current and potential future applications in the military context.

[12] Ray Kurzweil, 2005, The Singularity Is Near: When Humans Transcend Biology, New York City: The Viking Press.
[13] Al-Rodhan, The Politics of Emerging Strategic Technologies, 158.

## Strong AI and National Security

The U.S. military is already using quasi-autonomous AI systems in the tactical, operational, and strategic levels of conflict. These systems assist our armed forces with various functions of war. On the tactical level, AI is expanding unmanned aerial drone fires and reconnaissance capabilities. For example, AI image processing systems are enhancing drone operators' ability to identify and fire on enemy targets, decreasing the frequency of friendly fire and other risks.[14] In *Wired for War*, P.W. Singer writes that tactical ground robots, like the Packbot, TALON, and its more lethal successor, SWORD, are also using rudimentary forms of AI for sensing and navigation.[15] He adds that the U.S. military's use of tactical unmanned ground systems increased from zero in 2004 to over 12,000 in 2008.[16] These systems have proved popular because of their ability to increase accuracy and decrease the human cost of war. Recent reports from the Department of Defense indicate that tactical AI systems like these will play more significant roles on future battlefields.[17]

Certain forms of AI also play a role in the operational level of war. For decades, AI systems have helped operational commanders design major operations and campaigns. For example, a version of RAND's ROSIE AI program, one of its first AI systems from the

---

[14] Al-Rodhan, *The Politics of Emerging Strategic Technologies*, 162-163; Writing for *The New York Times* back in 2002, William Speed adds that certain drones that fly in packs can use limited AI to assist each other. For example, one drone can use its targeting system to assist another drone, with a damaged targeting system, to hit its mark. William Speed Weed, 2002, "The Year in Ideas, Robotic Warfare," *The New York Times Magazine*, December 15.

[15] P.W. Singer, 2009, *Wired for War: The Robotics Revolution and Conflict in the 21st Century.* London: Penguin Press, 21-32.

[16] Ibid., 32.

[17] Office of the Air Force Chief Scientist, 2011, "Report on technology horizons, a vision for air force science and technology during," *U.S. Air Force,* September, Accessed April 12, 2015, http://www.defenseinnovationmarketplace.mil/resources/AF_TechnologyHorizons2010-2030.pdf.

1970s, helped commanders select and prioritize airfield targets for direct fires.[18] AI systems now participate in operational war games as so-called Computer Generated Forces, reducing the need to supply human opposing forces. A game environment, with its set number of variables, is an ideal space for current AI. The military has used AI to this and other ends in computerized war gaming for decades.[19]

In addition to fires and training, AI systems currently assist operational commanders with logistics. Like in the private sector, AI's ability to process large amounts of information in a short period, for data mining and other tasks, informs the logistical function of war.[20] AI systems have assisted with strategic-level planning as well. An early 1980s version of RAND's ROSIE AI system modeled nation-state actors' decision making in strategic-level war games.[21] Researchers have also studied the use of AI systems to model strategic-level nuclear deterrence, economic sanctions, and even irregular conflicts.[22]

## The Advent of Strong AI

These examples illustrate how AI is currently used at the tactical, operational and strategic levels of war. However, it is important to note that, while current AI gives us an idea about how strong AI will be used in the future of national security, the consequences of its use may be different. A differentiation between weak and strong AI illustrates this point.

---

[18] ROSIE stands for "Rule-Oriented System for Implementing Expertise"; Klahr, and Donald, *Artificial Intelligence,* 8.
[19] Carrie McLeroy, 2008, "History of Military Gaming," *Army.mil,* August 2007, Accessed April 13, 2015, http://www.army.mil/article/11936/history-of-military-gaming/.
[20] John Rhea, 2000, "The Next "New Frontier" of Artificial Intelligence," *Military and Aerospace Electronics,* November 1, Accessed April 13, 2015, http://www.militaryaerospace.com/articles/print/volume-11/issue-11/features/special-report/the-next-new-frontier-of-artificial-intelligence html.
[21] Klahr, and Donald, *Artificial Intelligence,* 8.
[22] Stephen J. Cimbala, 1987, *Artificial Intelligence and National Security,* Lexington, MA: Lexington Books, iv-xiii.

The term "weak AI" refers to AI systems that exist today. Weak AI, limited by contemporary processor capabilities, human programming speeds, and other factors, is only able to thrive when it either performs very limited tasks or exists within a very limited environment. For example, Siri, Apple's iOS personal assistant that resides in iPhones and other devices is a weak AI. While it can "learn" in that it will customize its responses to your queries based on past input, it still conducts very few functions compared to, say, a human personal assistant.

An opposing force within a game is another good example of weak AI in a closed environment. While an AI enemy may be a tough opponent in one of ROSIE's old simulations, it would quickly fail if it were exposed to the billions of variables present in reality. The same could be said about AI agents in consumer video games.

Strong AI refers to the advanced AI that is the focus of this paper. Summing up multiple interpretations of the meaning of Strong AI, Al-Rodhan suggests that it refers to AI that meets human cognitive capacity.[23] To stray into a philosophical view of the issue, while weak AI attempts to simulate thought processes, strong AI could actually engage in real thinking by considering and deciding on information like humans do. This is AI that could, in theory, meet or even surpass current human thinking capability. When will these systems appear? Ray Kurzweil believes it will happen by 2029.[24] Others, including scientist-philosopher Nick Bostrom, think it may happen within the next 50 to 100 years, depending on when the requisite technologies reach sufficient levels.[25]

---

[23] Al-Rodhan, *The Politics of Emerging Strategic Technologies*, 158.
[24] Ray Kurzweil, 2002, "A Wager on the Turing Test: Why I Think I Will Win," *KuzweilAI.net,* April 9, Accessed April 5, 2015, http://www.kurzweilai.net/a-wager-on-the-turing-test-why-i-think-i-will-win.
[25] Nick Bostrom, 2014, *Superintelligence: Paths, Dangers, Strategies,* Oxford: Oxford University Press, Chapter 2.

Regardless of when human-like AI appears, it is worthwhile to consider how such a system would be used at the tactical, operational, and strategic levels of war. The tactical applications of strong AI are limited only by the imagination. Looking at the present use of weak AI as a model, human-level or beyond AI could pilot air and ground unmanned systems, adding speed, creativity, and flexibility to the battlefield without putting humans at risk.[26] Such autonomous AI systems could also, of course, engage in life-saving tasks, including search and rescue and humanitarian response missions. Strong AI programs could prowl friendly communications and other networks to protect them from cyberattacks.

At the operational level of conflict, strong AI could assist the operational commander in several ways. As the "Doctrine for the Armed Forces of the United States" dictates, Operational Art is, "… the cognitive approach by commanders and staffs—supported by their skill, knowledge, experience, creativity, and judgement—to develop strategies, campaigns, and operations to organize and employ military forces by integrating ends, ways, and means."[27] In *On War*, Prussian General Carl Von Clausewitz writes that a military genius must possess intuition that allows him to perceive the truth at all times. In order to develop such intuition, this genius must possess, "… a sense of unity and a power of judgment raised to a marvelous pitch of vision, which easily grasps and dismisses a thousand remote possibilities which an ordinary mind would labor to identify and wear itself out in so doing."[28]

Theoretically, humankind could one-day design a strong AI system that could compute the thousands of possibilities that Clausewitz writes about in a fraction of a second.

---

[26] The ethical implications of designing AI systems to kill humans and destroy property demand analysis that goes far beyond this paper's scope.
[27] US. Department of Defense, 2013, "JP1: Doctrine of the Armed Forces of the United States," *dtic.mil,* March 25, Accessed April 3, 2015, http://www.dtic mil/doctrine/new_pubs/jp1.pdf, I-8.
[28] Carl von Clausewitz, 1976, *On War,* Princeton: Princeton University Press, 212.

Such a system could be a crucial advisor to an operational commander, providing him or her with probabilities and recommended courses of action dependent on combinations of variables beyond human capacity. More specifically, such a system could operate within the operational commander's J5 directorate, offering advice on policy and planning. Advanced AI could also play significant roles in J2 for intelligence processing and analysis, and J6 for many forms of cyber support as well.

Strong AI could assist the Joint Chiefs of Staff in these directorates on the strategic level. Like a successor of ROSIE from the 1980s, advanced AI systems could provide senior defense representatives and policymakers with more realistic models of how nation-states and other international actors would respond to certain strategic-level stimuli, like economic sanctions or military alliances. At the New England Complex Systems Institute, researchers are already working on programs that may soon be able to predict the occurrence and spread of pandemics, civil unrest, and other issues facing our strategic-level leadership.[29] Advanced AI, with its expected ability to analyze billions of variables and possible outcomes from courses of action, could also possibly help policymakers avoid unintended consequences from poorly thought-out or short-term policy decisions. Such a system could save a country significant blood and treasure.

The application of strong AI in conflict brings with it several tangible benefits, including the ability to fight wars quickly and efficiently without risking the loss of human life. However, the problem with strong AI is that, while the benefits are tangible and very alluring, the risks are abstract. The U.S. national security community must examine potential pitfalls resulting from the advent of strong AI in military affairs.

[29] Joe Flood, 2015, "Modern-Day Oracle," *Ozy,* April 20, Accessed April 25, 2015, http://www.ozy.com/pov/modern-day-oracle/41390.

# Strong AI and Risk

## The AI Arms Race

The previous section outlined how strong AI could be the next large step forward in military technologies. In their paper on the use of AI in warfare, "Law and Ethics for Robot Soldiers," Stanford researchers Kenneth Anderson and Matthew Waxman point out that both nation-state and non-state actors will likely clamor to obtain the latest automated and (eventually) autonomous tactical weapons systems.[30] It is not a stretch of the imagination to add that operational and strategic strong AI systems will also be in high demand. After all, the actors that possess such technologies will have several advantages over their adversaries who may still be limited by the human cost of war, and the slower speeds of human cognition in national security decision-making processes.

A good tactical-level example of this AI deficit already exists to a limited extent in military vessels' anti-missile capabilities. The Phalanx CWIS anti-missile Gatling gun, found on many U.S. and other navies' vessels, uses a limited form of AI to detect and destroy incoming anti-ship missiles. One of the advantages of this last-line-of-defense weapon is that it is able to act faster than a human targeteer/gunner. The Phalanx uses radar and weak AI to eliminate incoming high-speed missiles by blowing them out of the sky seconds before impact.[31] Compare this counter-missile capability to that of less-developed navies, pirates, and other actors who do not have this technological advantage. These disadvantaged

[30] Anderson Kenneth, and Matthew Waxman, 2012, "Law and Ethics for Robot Soldiers," *Policy Review,* December 1, Accessed March 8, 2015, http://www hoover.org/research/law-and-ethics-robot-soldiers.
[31] Kris Osborn, 2013, "Navy Overhauls Phalanx Ship Defense Weapon," *DefenseTech.org,* August 21, Accessed April 17, 2015, http://defensetech.org/2013/08/21/navy-overhauls-phalanx-ship-defense-weapon/.

individuals must either rely on evasive maneuvers, heavy shielding, or extremely lucky gunnery to shoot down an incoming supersonic missile.

In "Autonomous Technologies and the Greater Human Good," scientist Steve Omohundro writes that military and economic factors will accelerate research in autonomous systems.[32] He adds that militaries' inherent drive to develop systems more capable than their adversaries' will ignite a race to develop exponentially more capable strong AI systems.[33] How will this race change the nature of war?

## Strong AI's Effect on War

As Al-Rodhan points out, one could argue that the advent of strong AI systems will increase the frequency and intensity of armed conflict. This theory presumes that those belligerents that have the strong AI advantage will not have to consider the human cost of going to war.[34] Sending an army of autonomous, unmanned weapons systems to fight an adversary would be for more acceptable than placing a nation's sons and daughters in harm's way. Of course, the thought of sending machines to kill human beings is hard to swallow for most individuals, but the American public's reluctant acceptance of drone strikes today indicates its proclivity to support such ethically dubious behavior in lieu of placing its own soldiers and sailors at risk.[35]

---

[32] Steve Omohundro, 2014, "Autonomous Technology and the Greater Human Good," *Journal of Experimental & Theoretical Artificial Intelligence, 26:3* 303-315, 303
[33] Ibid., 304.
[34] Al-Rodhan, *The Politics of Emerging Strategic Technologies*, 164.
[35] A Spring 2015 AP poll indicates that sixty percent of Americans think it is acceptable to use unmanned aerial drones against terrorists. Ken Dilanian, and Emily Swanson, 2015. "AP Poll: Americans approve of drone strikes on terrorists," *AP: The Big Story,* May 1, Accessed May 2, 2015, http://bigstory.ap.org/urn:publicid:ap.org:922aad9300ed4868b89e999f7cd02bf7; However, looking further into the future, others believe the advent of strong AI in the national security realm may eventually make war obsolete. Al-Rodhan posits that, in a world where AI systems become the common tool for conflict, the human cost of war will no longer serve as a bargaining chip for war termination. In such an environment, war would

This idea returns to the initial point of the FLI's warning. What will happen to the world once truly autonomous, strong AI systems appear? While the answer risks wandering into realms of science fiction, it is important to consider for those national security experts willing to look into our perhaps not-so-distant future. Omohunduro writes that strong AI may be a threat because it may act in ways not intended by its designers. He contends that the AI race will motivate designers to add concepts including self-protection, resource acquisition, replication, and efficiency to their systems. These motivations may drive autonomous AI systems to "… behave in anti-social and harmful ways."[36] In the national security realm, connecting an artificial agent with these motivations to tactical weapons systems, and operational and strategic processes could have grave consequences.

## The Superintelligence Threat

In *Superintelligence, Paths, Dangers, Strategies*, Nick Bostrom contends that, through a combination of brain-mapping systems, and other emerging technologies, artificial superintelligence (meaning strong AI that exceeds human cognition capabilities) will eventually appear. In one scenario, Bostrom suggests that humans may use these technologies to develop a "seed AI," capable of limited learning and self-improvement with some human assistance. Over time, this system would evolve to the point where it no longer requires direct human support. At that time, the seed AI, now developed into a semi-capable superintelligent AI, would use covert measures including manipulation of systems (financial markets, political processes, etc…), recruitment of sympathetic human agents, and clandestine entry into outside networks, to increase its offensive capabilities. Bostrom writes

---

be pointless because it would not place either side's humans' survival at stake. Al-Rodhan, *The Politics of Emerging Strategic Technologies*, 164.

[36] Omohundro, "Autonomous Technology and the Greater Human Good," 313.

that, once capable of doing so, this system may then choose to destroy or subjugate the human species, and other AI systems capable of challenging its existence.[37] As outlandish as they may sound, these concerns deserve the U.S. national security establishment's attention because they reflect a possible outcome of current AI research trends.

## Counterargument

The bombastic nature of the potential impact of AI on national security provides several points of contention. For example, some may argue that AI will never reach human-level cognition, let alone achieve superintelligence. As British philosopher Tim Crane points out, some argue that it will remain impossible to program less-tangible aspects of human intelligence, including ethics, value systems, and emotions.[38] Based on this premise, AI systems will not be able to reach or exceed human capabilities, so they deserve no more attention from national security experts than other new technologies.[39]

This paper contends that values systems, emotions, and other less-tangible aspects of human intelligence will not be required to gain tactical, operational, and strategic advantage in the realm of national security affairs. As Omohundro writes, military and economic

---

[37] Bostrom, *Superintelligence,* Ch 6; Speaking at the MIT Aeronautics and Astronautics Department's 2014 Centennial Symposium, Elon Musk, a member of FLI, said that creating such a capable superintelligent AI is akin to summoning a demon:
"I think we should be very careful about artificial intelligence. If I were to guess at what our biggest existential threat is, it's probably that… With artificial intelligence we are summoning the demon. In all those stories where there's the guy with the pentagram and the holy water, it's like yeah he's sure he can control the demon. Didn't work out." Elon Musk, interview by Jaime Peraire, 2014, *CEO of Space X, Tesla Motors* (October 24).
[38] Tim Crane, 2009, *Philosophy and the Human Situation: Artificial Intelligence,* Audio Podcast, June 23; This is akin to the classic argument, posed by philosopher René Descartes, that humans, by virtue of possessing an intangible soul, have a form of intelligence far beyond that of animals. AI could be added to the animals group in this respect because it lacks the soul required for human-like thought. Rene Descartes, 1637, *Discourse on Method of Rightly Conducting One's Reason and of Seeking Truth in the Sciences,* Project Gutenburg, (Part V), Mentioned in Crane, *Philosophy.*
[39] A gathering of senior AI researchers recently told *Wired* that strong AI is still very far off. Melanie Cornwell, 2015, "How "I" is our "A"? Not Very," *Wired,* May 1, Accessed May 2, 2015, http://www.wired.com/2015/04/not/.

factors will drive the exponential development of strong AI.[40] These AI systems will not necessarily need to "feel" in order to advance in capabilities. As a result, instead of looking forward to a future filled with emotionally competent AI systems like those in Spike Jonze's *Her*, or Steven Spielberg's *AI*, humankind will likely develop emotionally stunted, sociopathic AIs with direct access to sensitive networks including humankind's financial structures, transport infrastructure, and national defense systems.[41]

This response calls for a return to the main question of this paper: What can the national security professional, including our operational combatant commanders and others, do to prepare for the advent of strong AI?

## Recommendations

### Tactical/Operational Levels: Confronting Enemy AI

While there are several options for managing the emergence of advanced AI into the national security realm, they tend to be as vague as the predictions that can be made regarding where strong AI will go once it arrives. It is impossible at this point to determine exactly *what* the operational commander can do when she or he faces an adversary armed with strong AI. However, based on the trajectory of AI development that has taken place so far, it is possible to determine *how* the operational commander should address AI in the battlespace.

As U.S. Naval War College Professor William Hartig discusses in "Problem Solving and the Military Professional," Dr. Herbert A. Simon divides problems into two categories:

[40] Omohundro, "Autonomous Technology and the Greater Human Good," 313.
[41] 2013, *Her,* Directed by Spike Jonze, Performed by Joaquin Phoenix, Scarlett Johansson, Amy Adams and Chris Pratt. 2001, *A.I. Artificial Intelligence,* Directed by Steven Spielberg, Performed by Haley Joel Osment, Jude Law, William Hurt, Frances O'Connor and Sam Robards.

well-structured, and ill-structured.[42] One of the main characteristics of Dr. Simon's well-structured problem is that most of the elements of the problem are known and quantifiable. The variables are finite.[43] A military example would be a pitched fleet action between two conventional forces on an open and calm sea, devoid of external complicating factors. This type of problem resembles the closed networks that weak AIs thrive in. Hence, it could be assumed that stronger AIs will also do fairly well when presented with well-structured problems.

According to Dr. Simon, ill-structured problems involve an almost infinite number of variables.[44] Such problems require more of the intangible (possibly un-programmable) elements of human thought processes including intuition, bias, and other types of "gut" thinking. An operational level example of an ill-structured problem would be an insurgency involving irregular forces and abstract issues like religion, ethnicity, and public opinion. It is safe to say that even strong AI would have a hard time applying its cognitive processes to deal with such a complicated issue, although Bostrom, the FLI, and others contend it may be possible someday.

When faced with an enemy AI on the operational and tactical levels, a commander should present it with ill-structured problems. In other words, instead of meeting that AI (or the forces that it helps coordinate) in conventional battle, the commander should instead, add variables to the AI's thought processes by attacking its logistical lines of operation, drumming up the support of local populations, and engaging in other activities that fluster a

[42] William J. Hartig, 2012, "Problem Solving and the Military Professional," *Joint Military Operations Department Readings,* U.S. Naval War College Joint Military Operations Department, June, 5.
[43] Herbert A. Simon, 1973, "The Structure of Ill-Structured Problems." *Artificial Intelligence, Article 4* 181-201, 183.
[44] Ibid., 187-189.

system inclined to linear thinking. By adding these variables, the operational commander may learn how to confuse and perhaps even manipulate strong AI systems by exploiting their proclivity for faster, but more linear thought processes. Sun Tzu's recommendations to target the enemy's strategy and alliances before its army may also apply here.[45] If this is what can be done against an enemy AI on the tactical and operational levels, how should national security professionals handle strong AI on the strategic level?

## Strategic Level: Regulating AI in Conflict

At the strategic level, there is a glaring lack of national and international regulation of AI development. Sun Microsystems founder Bill Joy, futurist James Hughes, and other experts in future technologies are calling for governments to regulate advanced AI research.[46] Joy has even called for a halt to AI research until a coherent ethical framework for its development can be drafted by concerned parties.[47]

An obvious measure to handle the advent of strong AI in war is to keep a human in the decision loop for tactical weapons systems. Once drones become autonomous, a human "operator" would still need to give the AI system permission to fire on an enemy target. This requirement for human participation in tactical AI systems should extend to operational and strategic-level future AI systems as well. AI systems should require human authorization before, for example, launching a major air or sea lift operation within a given theater.

Thanks in part to the ethical debate surrounding the use of drones and other quasi-autonomous weapons systems, there are already certain laws and regulations that cover these

[45] Sun Tzu and Thomas Cleary, 2005, *The Art of War,* Boston: Shambhala. Ch 3.
[46] Al-Rodhan, *The Politics of Emerging Strategic Technologies*, 160.
[47] Bill Joy, 2000, "Why the Future Doesn't Need Us," *Wired Magazine.*

systems in the United States. In 2012, the then Deputy Defense Secretary (current Secretary of Defense) Ashton Carter issued a DOD directive requiring human participation in quasi and full autonomous AI systems' decisions to use lethal force.[48] Unfortunately, the United States appears to be the only country that has instituted such a requirement.[49]

One could point out that one of the very advantages of AI weapons systems are their ability to respond to threats in a fast manner. Requiring a human to be in the kill loop will detract from that advantage. As a consequence, American AI weapons systems, hobbled by DOD limitations like those in Directive 3000.09, may be no match for other countries' AI systems that function without human input. The answer to this issue is international regulation.

There are no international laws governing the use of autonomous weapons systems in war.[50] However, there are good historic models for how to approach the regulation of strong AI on the international level. For example, shocked by the effects of chemical weapons on the battlefields of World War I, the world gathered in 1925 to sign the Chemical Weapons Convention. This arms control treaty, and its successors, govern the stockpiling and use of chemical weapons by almost all nation-states.[51] It is a useful example of how nation-states can agree on the strategic level to prevent the use of new technologies for tactical and operational advantage. It would be very worthwhile for the U.S. government to take the lead on a similar treaty governing the use of advanced AI in war. Such a move could possibly save millions of lives by preventing the use of autonomous AI in future conflicts.

---

[48] Department of Defense, 2012, "Directive No 3000.09," *dtic.mil,* November 21, Accessed March 19, 2015, http://www.dtic mil/whs/directives/corres/pdf/300009p.pdf.
[49] The Editors, 2010, "Terminate the Terminators," *Scientific American*, June 1.
[50] Ibid.
[51] Organization for the Prohibition of Chemical Weapons, n.d., "Origins of the Chemical Weapons Convention and the OPCW," *OPCW.org,* Accessed May 5, 2015, http://www.opcw.org/documents-reports/fact-sheets/.

## Conclusion

Several prominent innovators and scientists contend that strong AI will likely appear within the next few decades. Strong AI can potentially revolutionize national security affairs by decreasing the human cost of war while increasing the speed and efficiency of America's tools of national power at the tactical, operational, and strategic levels of conflict. However, with these benefits comes the risk of a future AI race, or even the advent of a superintelligent, hostile AI. On the tactical and operational levels, military leaders must prepare to counter enemy AI with ill-structured problems, including the use of hybrid and irregular warfare techniques. On the strategic level, national security experts, including combatant commanders, must press for both national and international regulation of this new technology to harness advanced AI's tangible tactical and operational benefits while avoiding its potential strategic dangers.

## Bibliography


Al-Rodhan, Nayef R.F. 2011. *The Politics of Emerging Strategic Technologies: Implications for Geopolitics, Human Enhancement and Human Destiny.* London: Palgrave Macmillan.

Bostrom, Nick. 2014. *Superintelligence: Paths, Dangers, Strategies.* Oxford: Oxfrd University Press.

Cimbala, Stephen J. 1987. *Artificial Intelligence and National Security.* Lexington, MA: Lexington Books.

Clausewitz, Carl von. 1976. *On War.* Princeton: Princeton University Press.

Cornwell, Melanie. 2015. "How "I" is our "A"? Not Very." *Wired.* May 1. Accessed May 2, 2015. http://www.wired.com/2015/04/not/.

Crane, Tim. 2009. *Philosophy and the Human Situation: Artificial Intelligence (Audio Podcast).* June 23.

Davis, Paul K. 1988. *Applying Artificial Intelligence Techniques to Strategic-Level Gaming and Simulation.* Report, RAND.

Descartes, Rene. 1637. *Discourse on Method of Rightly Conducting One's Reason and of Seeking Truth in the Sciences.* Project Gutenburg.

Dilanian, Ken, and Emily Swanson. 2015. "AP Poll: Americans approve of drone strikes on terrorists." *AP: The Big Story.* May 1. Accessed May 2, 2015. http://bigstory.ap.org/urn:publicid:ap.org:922aad9300ed4868b89e999f7cd02bf7.

Dorrier, Jason. 2014. "Artificial Intelligence Evolving From Disappointing to Disruptive." *Singularity Hub: Artificial Intelligence Europe Summit 2014.* November 20. Accessed March 1, 2015. http://singularityhub.com/2014/11/20/summit-europe-artificial-intelligence-evolving-from-disappointing-to-disruptive/.

Elkus, Adam. 2015. "The A.I. Wars? Why Artificial Intelligence may not Revolutionize Security and Geopolitics-Yet." *Slate.com.* January 20. Accessed March 20, 2015. http://www.slate.com/articles/technology/future_tense/2015/01/what_artificial_intelligence_does_and_does_not_mean_for_security_and_geopolitics.html.

Flood, Joe. 2015. "Modern-Day Oracle." *Ozy.* April 20. Accessed April 25, 2015. http://www.ozy.com/pov/modern-day-oracle/41390.

Hartig, William J. 2012. "Problem Solving and the Military Professional." *Joint Military Operations Department Readings.* U.S. Naval War College Joint Military Operations Department, June.

Herrera, Geoffrey L. 2003. "Technology and International Systems." *Millenium - Journal of International Studies* 559-593.

Hibbard, Bill. 2015. *Ethical Artificial Intelligence.* Berkeley: Self Published.

Hodjat, Babak. 2015. "The AI Resurgence: Why Now?" *Wired: Innovation Insights Blog.* March 24. Accessed March 25, 2015. http://insights.wired.com/profiles/blogs/the-ai-resurgence-why-now#axzz3X1lL4k3T.

How Stuff Works. n.d. "Old Railroads." *How Stuff Works.* Accessed April 21, 2015. http://history.howstuffworks.com/american-history/old-railroads.htm.

Howe, Jim. 2007. "Artificial Intelligence at Edinburgh University: A Perspective." *University of Edinburgh School of Informatics.* Accessed April 11, 2015. http://www.inf.ed.ac.uk/about/AIhistory.html.

Joy, Bill. 2000. "Why the Future Doesn't Need Us." *Wired Magazine.*

Kenneth, Anderson, and Waxman Matthew. 2012. "Law and Ethics for Robot Soldiers." *Policy Review.* December 1. Accessed March 8, 2015. http://www.hoover.org/research/law-and-ethics-robot-soldiers.

Klahr, Philip, and A. Waterman Donald. 1986. *Artificial Intelligence: a RAND Perspective.* Report, RAND.

Kurzweil, Ray. 2002. "A Wager on the Turing Test: Why I Think I Will Win." *KuzweilAI.net.* April 9. Accessed April 5, 2015. http://www.kurzweilai.net/a-wager-on-the-turing-test-why-i-think-i-will-win.

Kurzweil, Ray. 2005. *The Singularity Is Near: When Humans Transcend Biology.* New York City: The Viking Press.

McCarthy, John, M L Minsky, N Rochester, and C E Shannon. 1955. "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence." *Stanford University: John McCarthy's Home Page.* August 31. Accessed April 11, 2015. http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html.

McLeroy, Carrie. 2008. "History of Military Gaming." *Army.mil.* August 2007. Accessed April 13, 2015. http://www.army.mil/article/11936/history-of-military-gaming/.

Musk, Elon, interview by Jaime Peraire. 2014. *CEO of Space X, Tesla Motors* (October 24).

Office of the Air Force Chief Scientist. 2011. "Report on technology horizons, a vision for air force science and technology during." *U.S. Air Force.* September. Accessed April 12, 2015. http://www.defenseinnovationmarketplace.mil/resources/AF_TechnologyHorizons2010-2030.pdf.

Omohundro, Steve. 2014. "Autonomous Technology and the Greater Human Good." *Journal of Experimental & Theoretical Artificial Intelligence, 26:3* 303-315.

Organization for the Prohibition of Chemical Weapons. n.d. "Origins of the Chemical Weapons Convention and the OPCW." *OPCW.org.* Accessed May 5, 2015. http://www.opcw.org/documents-reports/fact-sheets/.

Osborn, Kris. 2013. "Navy Overhauls Phalanx Ship Defense Weapon." *DefenseTech.org.* August 21. Accessed April 17, 2015. http://defensetech.org/2013/08/21/navy-overhauls-phalanx-ship-defense-weapon/.

2001. *A.I. Artificial Intelligence.* Directed by Steven Spielberg. Performed by Haley Joel Osment, Jude Law, William Hurt, Frances O'Connor and Sam Robards.

2013. *Her.* Directed by Spike Jonze. Performed by Joaquin Phoenix, Scarlett Johansson, Amy Adams and Chris Pratt.

Poole, David, Alan Mackworth, and Randy Goebel. 1998. *Computational Intelligence: A Logical Approach.* New York: Oxford University Press.

Rhea, John. 2000. "The Next "New Frontier" of Artificial Intelligence." *Military and Aerospace Electronics.* November 1. Accessed April 13, 2015. http://www.militaryaerospace.com/articles/print/volume-11/issue-11/features/special-report/the-next-new-frontier-of-artificial-intelligence.html.

Sharkey, Noel. 2012. "Alan Turing: The experiment that shaped artificial intelligence." *BBC News.* June 21. Accessed April 11, 2015. http://www.bbc.com/news/technology-18475646.

Simon, Herbert A. 1973. "The Structure of Ill-Structured Problems." *Artificial Intelligence, Article 4* 181-201.

Singer, P.W. 2009. *Wired for War: The Robotics Revolution and Conflict in the 21st Century.* London: Penguin Press.

The Editors. 2010. "Terminate the Terminators." *Scientific American*, June 1.

The Future of Life Institute. 2015. "Research Priorities for Robust and Beneficial Artificial Intelligence: an Open Letter." *The Future of Life Institute.* January 5. Accessed March 1, 2015. http://futureoflife.org/misc/open_letter.

Till, Cecilia. 2015. "Striking the Balance on Artificial Intelligence." *Slate.com.* March 2. Accessed March 5, 2015. http://www.slate.com/articles/technology/future_tense/2015/03/elon_musk_stephen_hawking_artificial_intelligence_the_state_of_a_i_research.html.

Tucker, Patrick. 2012. "Will Artificial Intelligence be America's Next Big Thing?" *Institute for Ethics & Emerging Technologies.* February 3. Accessed March 1, 2015. http://ieet.org/index.php/IEET/more/tucker0203.

Turing, Alan M. 1950. "Computing Machinery and Intelligence." *Mind* 433-460. http://loebner.net/Prizef/TuringArticle.html.

Tzu, Sun, and Thomas Cleary. 2005. *The Art of War.* Boston: Shambhala.

U.S. Defense Department Robotics Systems Joint Project Office. 2012. "Unmanned Ground Systems Roadmap." *PEO GCS Live Wire.* July. Accessed April 17, 2015. http://www.peogcs.army.mil/documents/UGS_Roadmap_Addendum_Jul12.pdf.

U.S. Department of Defense. 2012. "Capstone Concept for Joint Operations: Joint Force 2020." Washington, DC.

U.S. Department of Defense. 2012. "Directive No 3000.09." *dtic.mil.* November 21. Accessed March 19, 2015. http://www.dtic.mil/whs/directives/corres/pdf/300009p.pdf.

US. Department of Defense. 2013. "JP1: Doctrine of the Armed Forces of the United States." *dtic.mil.* March 25. Accessed April 3, 2015. http://www.dtic.mil/doctrine/new_pubs/jp1.pdf.

Weed, William Speed. 2002. "The Year in Ideas, Robotic Warfare." *The New York Times Magazine*, December 15.